Bhartiya Krishi Anusandhan Patrika, 36(3): 270-272 DOI: 10.18805/BKAP327

Advanced Strategies in Bulked Segregant Analysis and its Applications

Harshavardan J. Hilli

Department of Plant Breeding and Genetics,
Punjab Agricultural University, Ludhiana-141 027, Punjab, India.

Received: July 2021 Accepted: July 2021

ABSTRACT

Bulked segregant analysis (BSA) is a technique used to identify genetic markers associated with a mutant phenotype and is a quick method for identifying markers in particular genome regions. The paper focussed on Advanced methods which escape the requirement of genotyping all the individuals of the mapping population and generation of high-density linkage maps for mapping of the gene for the trait of interest. With the emergence of re-sequencing techniques, quick mapping of genes has become possible with reduced time and cost by using advanced methodologies like MutMap, MutMap+, MutMap-Gap, QTL-Seq, RNAseq BSA, NGS BSA and QTG seq. The procedure for various advanced BSA strategies has been described.

Key words: BSA, Mutmap, Mutmap+, Mutmap-gap, QTG seq, QTL seq.

INTRODUCTION

Bulked segregant analysis (BSA) is a technique used to identify genetic markers associated with a mutant phenotype. It is based on the principle of NILs and it was first described by Cliff Michelmore in Maize (1991).

Procedure for BSA

Forming two classes of opposite phenotypes for a trait of interest is part of the BSA technique. Bulked segregant analysis is a quick method for identifying markers in particular genome regions. The procedure compares two pooled DNA samples from individuals from a segregating population derived from a single cross. Individuals in each pool, or bulk, are similar for the trait or gene of interest but random for the rest of the genes. Two pools with opposing traits (for example, resistance and susceptible to a disease) are analysed to find markers that differentiate them. Polymorphic markers between the pools would be genetically linked to the loci that determine the trait used to construct the pools.

Terms related to BSA

Short reads

Short reads are short reads produced during genome sequencing that aid in examining the number of genomes derived from parents and distinguishing between them available SNPs.

SNP index

Index developed after aligning the data to the reference sequence of either parent. It ranges from 0 to 1. It is also defined as ratio of number of SNP's to the total reads.

Delta SNP index

Difference between SNP index of two bulks.

Bulked frequency ratio

Frequency of allele at each SNP position and ratio between bulks for each SNP.

Advanced stategies

Advanced methods escape the requirement of genotyping all the individuals of the mapping population and generation of high-density linkage maps for mapping of the gene for the trait of interest. With the emergence of re-sequencing techniques, quick mapping of genes has become possible with reduced time and cost by using advanced methodologies like MutMap, MutMap+, MutMap-Gap, QTL-Seq, RNAseq BSA, NGS BSA and QTG seq. Some of the examples of mapped genes using these methods mentioned in Table 1.

Mutmap

MutMap method was developed by Abe *et al* (2012)-mapping the mutation loci responsible for the change in leaf color from dark green to pale green in rice. The MutMap approach is based on a cross between the mutant and its wild form and thus specifically targets the causal SNPs that are responsible for phenotypic behaviour during mutagenesis.

A stable mutant will be crossed with wild-type parent to generate F2 population. F2's DNA Plants with the recessive mutant phenotype will be bulked, sequenced, matched to the reference genome and a sliding window method can be used to measure the SNP index. The causal SNP for the trait of interest will have an SNP index near 1, while the rest of the loci with a low SNP index depicts non-causal SNPs. (Kishore *et al*., 2018).

Mutmap +

When the development of an F2 mapping population is impossible due to lethality or sterility of the mutant triggered

*Corresponding author's E-mail:harshajh1995@gmail.com

Table 1: Examples of mapped genes.

Strategies	Crop	Mapped gene	Trait
Mutmap	Rice	CA01	Leaf color
Mutmap+	Rice	OsLAP6	Sterility
Mutmap gap	Rice	Pi	Blast resistance
RNA Seq	Wheat	Yr15	Rust resistance
NGS BSA	Soybean	qCC1 and qCC2	Cotyledon color

by recessive mutations, the mutmap strategy is used. The MutMap methodology was developed by Fekih *et al.* (2013) in rice.

In the MutMap+ process, 20-30 plants with mutant and wild-type phenotypes will be chosen in M3 generation to generate wild-type and mutant bulk for sequencing, respectively. Identification of SNPs for both wild type bulk as well as mutant bulk will be used for comparing both bulk sequences independently with reference genome and then the SNP index for wild-type bulk as well as mutant bulk will be calculated. By subtracting the wild type SNP index from the mutant SNP index, the Delta SNP index can be known. Positive SNPs index values indicate the presence of causal SNPs that cause phenotype.

Mutant population development and mapping in Mutmap + is similar to MutMap except for the following differences:-

o (1) No genetic hybridization is involved.

o (2) Progeny testing is mandatory as the surviving heterozygous plants of the wild type can be differentiated from the homozygous plants only through progeny testing.

Mutmap gap

MutMap-Gap is a MutMap extension specifically designed for identifying causative mutations in the reference genome's gap area. Demonstrated Mut- Map-Gap methodology in *Oryza sativa* cv. Hitomebore by mapping the gene responsible for rice blast resistance (Pii), which is located in the gap region of reference genome *O. sativa* cv Nipponbare (Takagi *et al.* 2015).

Eg:- (Takagis experiment)

√ In the first step of mapping, Hitomebore genome sequence was aligned on the Nipponbare reference genome replacing Nipponbare nucleotide with those of Hitomebore.

√ By inoculating Hitomebore mutants with Magnaporthe oryzae isolate TH68-126, a total of 3033 EMS-generated mutants were screened for blast resistance. By selfing mutant plants, the blast susceptible mutant line Hit5948 was discovered and made M2 advance generation.

√ The M2 blast prone mutant was crossed with the wild type parent, resulting in F1 and F2 populations. To produce mutant type bulk, 20 mutant plants were chosen from the F2 population.

√ The obtained sequence reads from mutant bulks was mapped on Hitomebore reference genome and used to search for the causative SNPs having SNP index * 1.

√ However, in the assembled region, SNPs with SNP index 1 were undetected, increasing the risk that the causative SNPs are in the gap region. The causal mutation in de novo assembled sequences was detected using the SNP index analysis, which was applied to all unassembled reads. Four SNPs with SNP index 1 were discovered in the gene Os09t0327600-01, which is located in the gap field, according to SNP index study.

QTL Seq

• QTL-Seq is an extension of BSA and MutMap methodology for mapping major effect quantitative trait loci (QTL) by using NGS platform.

QTL-Seq is a hybrid of BSA and MutMap, with the goal of mapping major effect quantitative trait loci (QTL) using a next-generation sequencing (NGS) platform (Pandey *et al.*, 2017). Existing BSA and MutMap methods can only discuss qualitative traits regulated by a single gene that result in distinct phenotype groups. In QTL-Seq, 10-20 individuals representing extreme phenotypes will be selected for preparing two bulks that is 'highest' bulk and 'lowest' bulk. Each bulk DNA will be sequenced and aligned independently to reference genome to calculate the SNP index. SNP index of lowest bulk is subtracted from SNP index of highest bulk, this result in formation of a peak in the plot of SNP index giving the position of the QTL.

QTG seq

QTG seq is a method to accelerate QTL fine mapping that combines QTL partitioning (to convert quantitative trait in to near qualitative trait) and sequencing of bulked pool from large segregating population (Zhang *et al.*, 2019).

The QTG-seq strategy consists of the following steps:-

• The populations F1, F2 and BC1F1 are used. If the researchers are unsure about the QTL mapping findings by F2 population to assess the location of the QTL, F2:3 families can be used.

• QTL partitioning, in which BC1F1 plants heterozygous for the QTL of interest but homozygous for the other QTLs will be selected and the desired BC1F1 plants will be self-pollinated to produce BC1F1:2 families. The derivative BC1F1:2 families will further be evaluated for the given trait and the bottom and top 20% of plants showing extreme phenotypes will be selected separately within each family.

• DNA from the two groups of selected plants will be extracted and bulked in equal amounts, forming two pools described as the "high pool" and "low pool".

• Whole-genome sequencing with deep genome coverage will be performed on the two wide DNA pools. The sequence data will be then compared to the reference genome and genetic variations can thus be detected around the entire genome.

• For each pool, the allele frequency will be determined. To reliably evaluate the location of the target locus, a new statistic called smooth LOD, the smoothed version of the LOD ranking, can be suggested to assess the target locus's peak location.

Applications of advanced BSA strategies

√ Helps to identify markers associated with mutant phenotype.
√ Identification of mutation located in the gap region of the reference genome.
√ Helps to identify genes with quantitative effect even when crossing cant be done between wild and mutant.
√ Mapping simultaneously two qualitative genes.
√ For effective MAS *i.e* large population can be grouped in to sub populations based on MTA and then using BSA for improving efficiency of selection for sub population.
√ RNA seq helps to map genes even in populations for which no polymorphic markers have been previously identified.

REFERENCES

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M. and Innan, H., (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. Nature Biotechnology. 30(2): 174.

Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., Sharma, S., Sharma, S., Kanzaki, H., Matsumura, H. and Saitoh, H. (2013). MutMap+: Genetic mapping and mutant identification without crossing in rice. PloS One. 8(7): 180.

Kishore, U., Kumar, K., Sevanthi, A.M. and Gaikwad, K., (2018). MutMap: A versatile tool for identification of mutant loci and mapping of genes. Indian Journal of Plant Physiology. 23(4): 612-621.

Pandey, M.K., Khan, A.W., Singh, V.K., Vishwakarma, M.K., Shasidhar, Y., Kumar, V., Garg, V., Bhat, R.S., Chitikineni, A., Janila, P. and Guo, B. (2017). QTL seq approach identified genomic regions and diagnostic markers for rust and late leaf spot resistance in groundnut (*Arachis hypogaea* L.). Plant Biotechnology Journal. 15(8): 927-941.

Takagi, H., Tamiru, M., Abe, A., Yoshida, K., Uemura, A., Yaegashi, H., Obara, T., Oikawa, K., Utsushi, H., Kanzaki, E. and Mitsuoka, C., (2015). MutMap accelerates breeding of a salt-tolerant rice cultivar. Nature Biotechnology. 33(5): 445-449.

Zhang, H., Wang, X., Pan, Q., Li, P., Liu, Y., Lu, X., Zhong, W., Li, M., Han, L., Li, J. and Wang, P., (2019). QTG-Seq accelerates QTL fine mapping through QTL partitioning and whole-genome sequencing of bulked segregant samples. Molecular Plant. 12(3): 426-437.